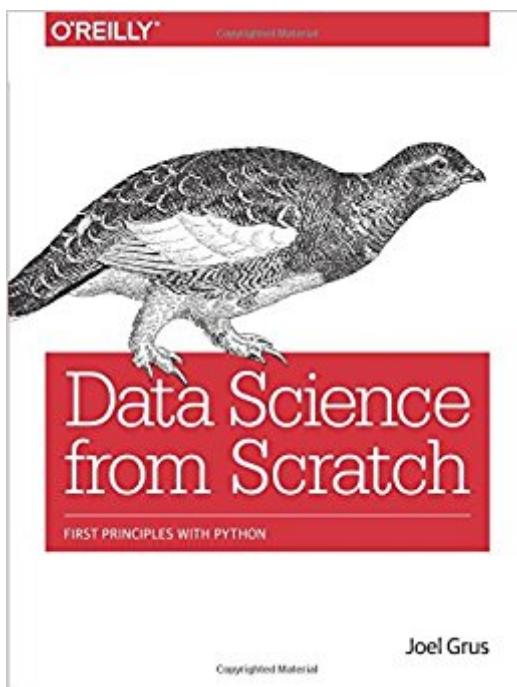The book was found

Data Science From Scratch: First Principles With Python

Synopsis

Data science libraries, frameworks, modules, and toolkits are great for doing data science, but theyâ ™re also a good way to dive into the discipline without actually understanding data science. In this book, youâ ™ll learn how many of the most fundamental data science tools and algorithms work by implementing them from scratch.If you have an aptitude for mathematics and some programming skills, author Joel Grus will help you get comfortable with the math and statistics at the core of data science, and with hacking skills you need to get started as a data scientist. Todayâ ™s messy glut of data holds answers to questions no oneâ ™s even thought to ask. This book provides you with the know-how to dig those answers out.Get a crash course in PythonLearn the basics of linear algebra, statistics, and probabilityâ ”and understand how and when they're used in data scienceCollect, explore, clean, munge, and manipulate dataDive into the fundamentals of machine learningImplement models such as k-nearest Neighbors, Naive Bayes, linear and logistic regression, decision trees, neural networks, and clusteringExplore recommender systems, natural language processing, network analysis, MapReduce, and databases

Book Information

Paperback: 330 pages

Publisher: O'Reilly Media; 1 edition (April 30, 2015)

Language: English

ISBN-10: 149190142X

ISBN-13: 978-1491901427

Product Dimensions: 7 x 0.7 x 9.2 inches

Shipping Weight: 1.1 pounds (View shipping rates and policies)

Average Customer Review: 3.9 out of 5 stars   88 customer reviews

Best Sellers Rank: #12,043 in Books (See Top 100 in Books)   #1 inÂ Books > Computers & Technology > Programming > Software Design, Testing & Engineering > Structured Design   #3 inÂ Books > Computers & Technology > Networking & Cloud Computing > Data in the Enterprise #7 inÂ Books > Computers & Technology > Databases & Big Data > Data Modeling & Design

Customer Reviews

Joel Grus is a software engineer at Google. Before that he worked as a data scientist at multiple startups. He lives in Seattle, where he regularly attends data science happy hours. He blogs infrequently at joelgrus.com.

This is a great book-- well written, easy to digest and informative. I've been in Data Mining and Statistical Analysis for a little over a decade now; I was looking for a book to share with my team to ensure we were all up-to-speed on some foundational concepts: this book is it. EDIT: I also forgot to mention, it has probably the best get-up-and-running in Python introduction I've seen (see, e.g., Chapter 2, ~20pp.)It's the right size and correct coverage for the content and the author's sense of humor (indeed, that of a data scientist) resonates with the audience.Solid introduction, even better review or brief explanation of commonly encountered topics.One of the best O'Reilly books I've read in a long time-- in fact, a technical book at the level I used to expect from O'Reilly.

Great book for those interested in learning or refreshing their memory on data science using python. The book takes you through different examples of topics and the python code you would use. I did find, however, that the descriptions and mathematics behind some of the concepts were a little lacking but still great in tandem with other resources.

It's a great start if you want to understand data sciences. Another thing this book does particularly well, is that it tell you what the next step are in term of other books to reads. I have not try using the codes example yet but I ll do so in the future. That s being said if you are just curious about data science and you are new to the field, you can skip the code part. I love this book.

Minus one star for using outdated Python 2.7. Essentially ALL data science tools you are likely to run across have been updated to Python 3.4+. I would have knocked off two stars but this book is actually quite good and delivers on its title.This is a very basic book on Data Science but it gives a broad overview which helps you get a perspective on the tools that are available. This book teaches methods by developing actual code for these methods. You will find in work situations that you will use library functions instead of "rolling your own" but this book helps bring the details together by having you actually code these techniques. I support this approach 100% Once you have this overview, you can drill down into specifics with other materials like textbooks or cookbooks.I'd did flinch at some of the explanations in this book but it really is a "from Scratch" approach and some things are simplified to avoid distractions.This book also teaches basic Python 2.7 with a quick start chapter, so it is self contained for any scientist or engineer that wants to get started adding Data Science techniques to their repertoire.

This is among the handful of very best technical books I have ever read.As the "from Scratch" in the

title implies, the objective of this book is to teach the fundamental ideas and techniques of data science from first (or nearly first) principles. After working through this book, you'll be better able to meaningfully utilize the pre-packaged software (whether it's Matlab, R, scikit-learn, or whatever) that you will use in "real life".And although the knowledge you'll gain is largely independent of the programming language, you will as a bonus learn from the clear and elegant python code included. Every key topic, from probability, statistics, and other mathematical subjects, to machine learning and databases, is covered in a crystal clear manner.In summary, this book is the bee's knees.

Its very good book if you have decent knowledge of stats/data science and need to refresh your skill or for R user who quickly want to learn things in python. For beginner without any stat and math background it is going to be Data science from scratching your head instead.

In my view, too many people want to be data scientist and use advance statistical techniques that they don't fully understand. I believe Grus closes this gap with this fun introductory book to some of the basic techniques in data analysis. If you don't come from a statistical background (MS/PhD), then you should start with this book and move on to other texts such as Pattern Recognition and Neural Networks by Brian D. Ripley for a thorough review of classification methods, or the more popular and modern The Elements of Statistical Learning By Hastie, Tibshirani and Friedman.

Great book, good first principal understanding of core data science concepts. Code along with the examples as you will need the functions you write throughout the book.

Download to continue reading...

Python: The Complete Python Quickstart Guide (For Beginner's) (Python, Python Programming, Python for Dummies, Python for Beginners) Python: Programming: Your Step By Step Guide To Easily Learn Python in 7 Days (Python for Beginners, Python Programming for Beginners, Learn Python, Python Language) Hacking with Python: Beginner's Guide to Ethical Hacking, Basic Security, Penetration Testing, and Python Hacking (Python Programming, Hacking, Python Coding, Python and Hacking Book 3) PYTHON: PYTHON'S COMPANION, A STEP BY STEP GUIDE FOR BEGINNERS TO START CODING TODAY! (INCLUDES A 6 PAGE PRINTABLE CHEAT SHEET)(PYTHON FOR BEGINNERS, PYTHON FOR DUMMIES, PYTHON PROGRAMMING) Data Analytics and Python Programming: 2 Bundle Manuscript: Beginners Guide to Learn Data Analytics, Predictive Analytics and Data Science with Python Programming Python Programming: Python Programming for Beginners, Python Programming for Intermediates, Python Programming

for Advanced PYTHON: LEARN PYTHON in A Day and MASTER IT WELL. The Only Essential Book You Need To Start Programming in Python Now. Hands On Challenges INCLUDED! (Programming for Beginners, Python) Data Science from Scratch: First Principles with Python Data Analytics: What Every Business Must Know About Big Data And Data Science (Data Analytics for Business, Predictive Analysis, Big Data Book 1) Data Analytics: Applicable Data Analysis to Advance Any Business Using the Power of Data Driven Analytics (Big Data Analytics, Data Science, Business Intelligence Book 6) Big Data For Business: Your Comprehensive Guide to Understand Data Science, Data Analytics and Data Mining to Boost More Growth and Improve Business - Data Analytics Book, Series 2 Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) Python: Learn Python in a Day and Master It Well: The Only Essential Book You Need to Start Programming in Python Now Maya Python for Games and Film: A Complete Reference for Maya Python and the Maya Python API Python Programming: An In-Depth Guide Into The Essentials Of Python Programming (Included: 30+ Exercises To Master Python in No Time!) Python: The Fundamentals Of Python Programming: A Complete Beginners Guide To Python Mastery. Python Data Science Handbook: Essential Tools for Working with Data Unsupervised Machine Learning in Python: How to Find Distinct Patterns in Your Data Without Being at the Mercy of Data Labeling by Third-Party Workers Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data (Princeton Series in Modern Observational Astronomy) Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython